The Thesis Committee for Lindsey Nan Wolf
Certifies that this is the approved version of the following thesis:

Characterization of beneficial mutations in unsaturated fatty acid biosynthesis that are recurrent dead-ends in a long-term evolution experiment with *Escherichia coli*

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor:

Jeffrey E. Barrick

James J. Bull

Characterization of beneficial mutations in unsaturated fatty acid biosynthesis that are recurrent dead-ends in a long-term evolution experiment with *Escherichia coli*

by

Lindsey Nan Wolf, B.A.

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

May 2014

Dedication

This work is dedicated to my family. I'm deeply grateful for their boundless support and love throughout this time. I also want to thank my friends who supply vast patience and enthusiasm and keep me going through all trials and successes.

Acknowledgements

Most importantly I want to thank Jeffrey Barrick for his guidance, encouragement, and dedication. He has become an excellent mentor and I am grateful to have had the experience of watching the Barrick lab grow and flourish since its inception. I also want to thank all the members of the Barrick lab who's enthusiasm for science is infectious and friendship is invaluable. This work was supported by the U.S. National Institutes of Health (R00-GM087550) and U.S. Army Research Office (W911NF-12-1-0390).

Abstract

Characterization of beneficial mutations in unsaturated fatty acid biosynthesis that are recurrent dead-ends in a long-term evolution experiment with *Escherichia coli*

Lindsey Nan Wolf, B.A.

The University of Texas at Austin, 2014

Supervisor: Jeffrey E. Barrick

Microbes provide an invaluable tool for watching evolution in action. Throughout more than 55,000 generations, lineages of *Escherichia coli* cells in a long-term evolution experiment (LTEE) grew in a minimal glucose environment and explored different mutational paths to higher fitness. Genome sequencing identifies genes that accrue mutations early in evolution across the twelve evolving populations. These parallel mutations typically provide a significant fitness benefit and often fix in the population. However, some mutations seem to lead to evolutionary dead-ends. In 7 of the 12 LTEE populations, lineages with mutations in the gene coding for the lipid synthesis repressor, *fabR*, gain traction within the population, but always eventually go extinct. To parse out the fitness benefits and downstream effects, strains with these mutations were constructed. These mutations increase the growth rate and may affect the length of lag phase after each daily transfer. Another mutation that often fixes within eventually

successful clades is within the stress response global regulator *spoT*. A connection between *spoT* and *fabR* mutations could be the key to understanding the eventual outcomes within these lineages. Decreased fatty acid synthesis (repressed by FabR) during glucose starvation activates the global repressor SpoT to produce the cellular "alarmone" (p)ppGpp, inhibiting cell growth during the stringent response. Thus, it is possible that *fabR* mutations that prolong fatty acid synthesis and *spoT* mutations that alter the production of (p)ppGpp may both benefit cells by affecting the stringent response. In addition, when these two mutations are combined in a single strain they confer nearly an identical increase in fitness as the single mutations alone, strengthening the argument that they may target the same cellular pathway. Preliminary gene expression analyses of *fabR* mutants confirmed an expected increase in unsaturated fatty acid synthesis and also found signs that membrane damage responses were activated. It is possible that *fabR* mutants are near a stability cliff that makes them unable to access otherwise beneficial further mutations. Ultimately, this work will elucidate how interactions between the physiological effects of mutations on evolutionary paths to higher fitness may lead to differences in evolvability that ultimately determine success or extinction.

Table of Contents

List of Tables

List of Figures

I. INTRODUCTION

EVOLUTION IN MICROBES

Understanding evolutionary processes is often limited by the practical difficulties of watching evolution as it occurs. Because of this, many experimental evolutionary biologists turn to microbial organisms to overcome these limitations.[1,2] Microbial populations are an ideal model system because they replicate rapidly, the conditions on which they grow can be precisely controlled, and they are more genetically tractable than most other organisms. So these experimental systems allow us to strip down evolution to a simple model form, witness it as it occurs, and begin to understand the molecular mechanisms driving it forward.

In the simplest model of microbial evolution, populations are completely asexual and begin as clones derived from single cells. This means that genetic changes are introduced into a lineage exclusively due to spontaneous mutations in the chromosome. Mutations that increase fitness may eventually sweep through the population and fix so that all other lineages are outcompeted and go extinct. However, large asexual populations often simultaneously contain multiple lineages following different mutational paths towards higher fitness (Fig. 1).[3,4] These subpopulations can coexist for a while, competing for available resources until one gains new, better mutations and takes over the population. These dynamics can result in situations where several lineages with varying fitnesses coexist long enough that the one that accrues the highest fitness after later mutations accumulate may not have been the most fit initially.[5,6]

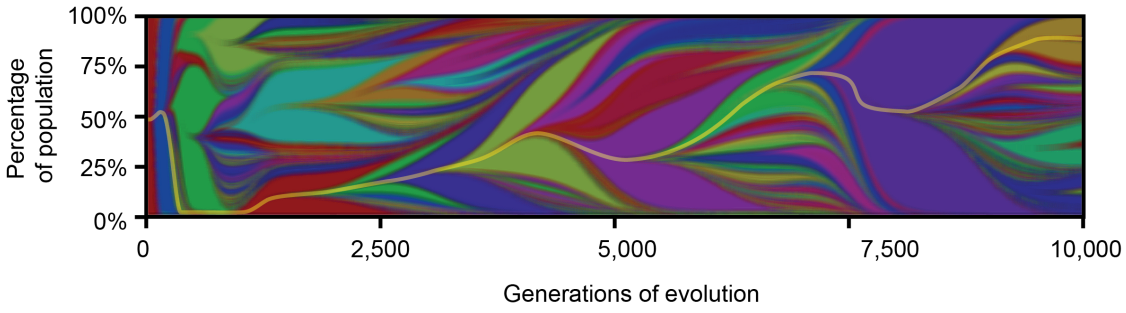

Figure 1. Evolutionary dynamics in a simulated *E. coli* population.

In this Muller plot each color wedge represents a new beneficial mutation arising in its parental type. The yellow line traces mutations in the lineage that dominates at 10,000 generations in the simulation.[13]

Perhaps the best-known microbial evolution experiment is the Lenski long-term evolution experiment (LTEE).[7–9] In the LTEE, two variants of a strain of *E. coli* B, differing by a neutral marker, were inoculated into 12 flasks containing minimal glucose media. Each day 1% of each culture was transferred into a new flask.[10,11] These populations have continued to be transferred for more than 25 years, allowing over 55,000 generations of bacterial history to accumulate. Every 500 generations, glycerol stocks were frozen for each population. Frozen samples could then be revived and competed against the ancestor to measure fitness advantages gained by the evolved populations.[12] Additionally, as technology has advanced since the beginning of the LTEE, the genomes of evolved strains can now be sequenced to comprehensively identify the mutations responsible for adaptation.[13–15]

EVOLUTIONARY DEAD-ENDS IN BACTERIAL POPULATIONS

In one of the 12 populations of the LTEE, named Ara-1, a number of early mutations were identified and their timing and order were reconstructed.[16] As in other microbial evolution experiments, multiple coexisting subpopulations vied for dominance before the "winners" eventual took over the population. Surprisingly, after isolating strains from the 500-generation subpopulations whose descendants eventually won and lost, it was found that the losers had a competitive advantage over the winners at this early time point (Fig. 2).

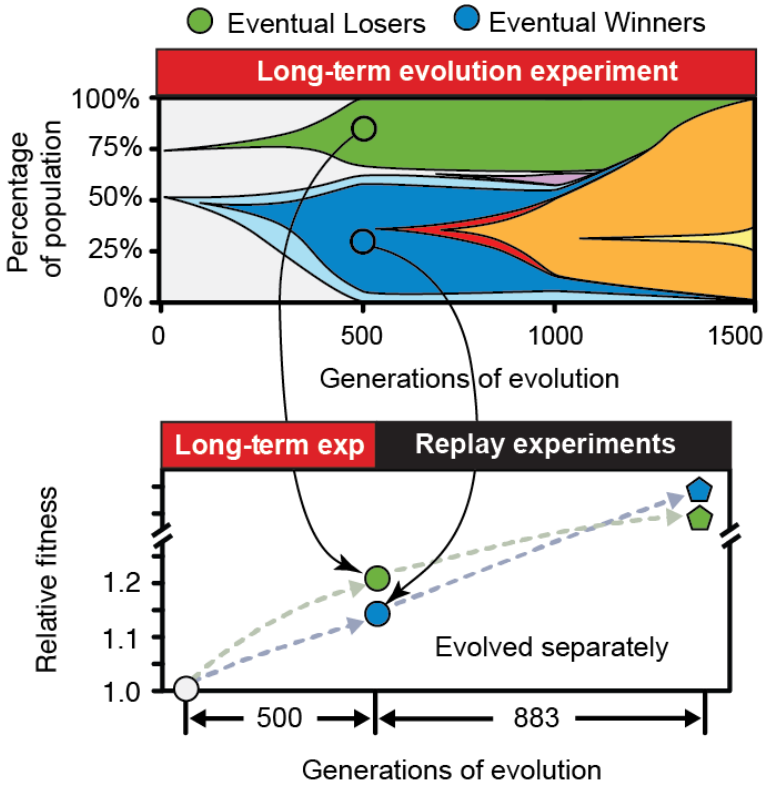

Figure 2. Fitness trajectories of eventual winner and loser types of *E. coli*.

Each was isolated from the long-term population after 500 generations. Then, they were evolved separately in replicate "replay" evolution experiments for 883 generations. Despite initially being behind at 500 generations in the original population, the eventual winner types reproducibly evolved to higher fitness.[8,16]

Moreover, in replay evolution experiments, using these 500-generation clones as a starting point, descendants of the "winners" were able to achieve greater fitness than descendants of the "losers" after an interval of further evolution. Thus, some early mutations in these subpopulations were affecting their ability to evolve, and it hadn't been a random event that they had caught up and outpaced the eventual losers in the original population. Woods et al. focused on two different mutations in the topoisomerase gene *topA* characteristic of the eventual winners (EW) and eventual losers (EL) and how these mutations interacted with a later mutation in *spoT* to result in the inability for the ELs to benefit from a *spoT* mutation in the same way that the EWs did (Fig. 3).

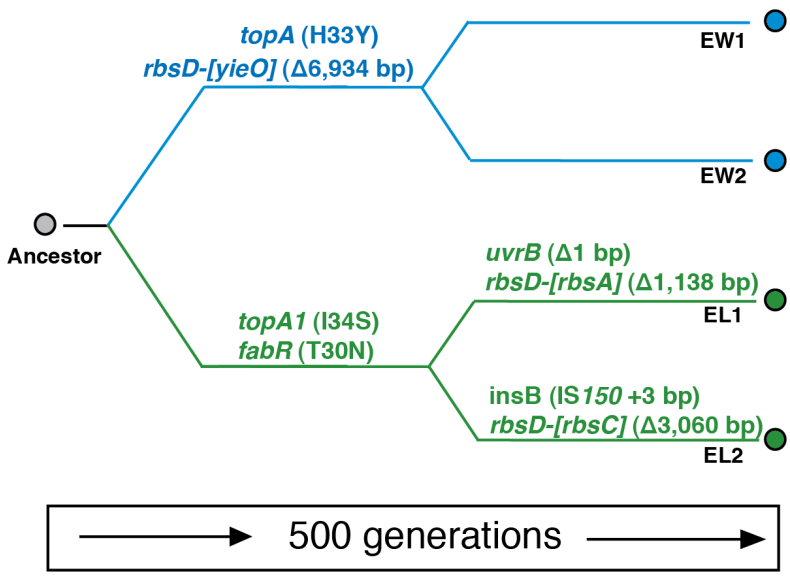

Figure 3. Distinct subpopulations occur in the same flask.

Mutations labeled on branches were identified through whole-genome resequencing of evolved eventual winner and eventual loser clones from the endpoint of the replay experiment.[16]

MUTATIONS IN REGULATORY GENES AFFECT GROWTH

One mutation that existed early in the eventual loser subpopulation, but had not been a focus for Woods et al. is in a gene for the unsaturated fatty acid (FA) synthesis repressor, *fabR*. This repressor blocks the expression of the FA synthesis enzymes FabB and FabA.[17] FabR is active during nutrient starvation, when it dimerizes and binds the promoter regions of *fabB* and *fabA*. In addition, free acyl carrier protein (ACP) is a required cofactor for FA synthesis. When ACP is free or covalently attached to saturated fatty acid chains it is unable to facilitate FabR repression so *fabA* and *fabB* are expressed and fatty acids are synthesized. However, when ACP is attached to unsaturated fatty acid chains it enables FabR-mediated repression. In this way ACP "senses" the need for unsaturated fatty acid synthesis and provides a feedback mechanism when FA synthesis is unnecessary.[18] Downstream effects of FabR repression can be far-reaching. A mutation in this gene and subsequent continuation of FA synthesis would likely delay normal processes that FabR repression would initiate.

One downstream system affected by a decrease in FA synthesis is the stringent response. This is a cellular process initiated in low nutrient environments by either amino acid starvation, activating RelA, or FA synthesis inhibition, activating SpoT.[19,20] Both RelA and SpoT are global regulatory proteins. Each detects nutrient starvation and begins to synthesize guanosine 5'-(tri)diphosphate, 3' diphosphate [(p)ppGpp] from GTP. (p)ppGpp is an "alarmone" that results in expression of cellular stress resistance genes

and growth arrest.[19] Interestingly, one of the mutations that occurs later in the eventual winners is in *spoT*. SpoT can both synthesize and degrade (p)ppGpp, depending on environmental conditions. During FA starvation, binding to unacylated-ACP modulates the activity of SpoT so that it synthesizes (p)ppGpp and triggers the stringent response. Since acylated ACP is a FA precursor molecule, the presence of free ACP during FA repression can act as an indicator of FA starvation to SpoT and activate (p)ppGpp synthesis and the resulting stress response.

Mutations in both *fabR* and *spoT,* blocking FA repression or (p)ppGpp synthesis, respectively, could delay or alter the threshold with which the stringent response is activated. Hypothetically, in the LTEE this could be beneficial to cells because they could continue dividing for longer as glucose is depleted or shorten their lag phase when coming out of stationary phase during each daily cycle of serial transfer to fresh media. My goals were to characterize competing subpopulations early in the other eleven LTEE populations to find other cases of similar EW/EL dynamics and to elucidate the roles of early mutations in *fabR* and other global regulatory genes and how their effects interact within the cell to allow continued evolution or, as in the case of the eventual losers, why they seem to repeatedly lead early clades down a longer-term dead-end.

II. MATERIALS AND METHODS

LONG-TERM EVOLUTION EXPERIMENT

Samples were obtained from Richard Lenski (Michigan State University).

RIBOSE DELETION GENOTYPING

In two populations, Ara-3 and Ara-5 at 500 and 1000 generations, 95 clones were randomly picked and genotyped by whole-cell PCR to detect fragment length differences indicative of IS*150*-mediated deletions overlapping the ribose operon using *rbs* primers (1, 19, and 20) (Table 1). Selected isolates with different deletion sizes were Sanger sequenced to determine the precise chromosomal bases deleted. These clones were frozen and saved for later identification of secondary beneficial mutations.

To determine which secondary mutations occurred in each lineage, up to three (if possible) representative clones were selected for each unique ribose operon deletion. Candidate genes, such as *nadR, pykF, hslU*, and *infB* were then amplified from the genome via PCR and Sanger sequenced (49-56 and 119-127 in Table 1). For both populations, Ara-3 and Ara-5, I profiled *rbs* deletions at 500 and 1000 generations. Since eventual loser *fabR* mutants are present later in Ara-3, several clones with each ribose deletion were analyzed in samples through 4000 generations in this population.

WHOLE-GENOME RESEQUENCING

Samples of all twelve LTEE populations archived at 500 and 1000 generations were revived to isolate metagenomic DNA for next-generation sequencing. To minimize the possibility that regrowth distorted the true allele frequencies present within the

original population when it was frozen, we washed cells twice in DM0 media to remove the glycerol used as cryoprotectant (which can also be a carbon source for *E. coli*) and initiated 10 ml DM25 cultures from 120 μl of each frozen stock to maintain approximately the normal number of cells that would be present after a normal daily transfer in the LTEE.

Mixed-population DNA was purified from these cultures after 24 h of growth using the Invitrogen PureLink Genomic DNA kit, then prepared for Illumina sequencing using standard procedures for fragmenting and ligating bar-coded DNA adaptors. Illumina HiSeq sequencing at the University of Texas at Austin Genome Sequencing and Analysis Facility (GSAF) yielded 100-base paired-end reads for each sample. Reads were analyzed using the *breseq* computational pipeline (version 0.24) in polymorphism mode using the genome of the LTEE ancestral strain REL606 as a reference sequence. Each sample had >100-fold read depth coverage. The frequencies of *fabR*, *rbs, spoT*, and *topA* mutations were tabulated from this data.

CONSTRUCTING MUTANT STRAINS

In order to construct a strain with a *fabR* mutant allele in an ancestral background, I used the allelic exchange plasmid pKOV to generate an unmarked mutation as previously described.[21] Although there are a number of mutations that occur within this gene, I chose the point mutation found in the eventual loser lineage in Ara-1 (T30N). This allele was amplified via PCR using primers (#108, #109) from the genome of the 500-generation Ara-1 strain TC1135. Allelic exchange vector pKOV with the EL-*fabR*

allele was then constructed and transformed into the ancestor strain REL606. Double mutants (*spoT* and *fabR*) were constructed in the same way, using the *fabR* mutant pKOV construct and transforming that into previously constructed strain TC710,[22] with a *spoT* mutant allele in an ancestral background. The *fabR* mutant and the *spoT* and *fabR* double mutant were designated LNW24 and LNW26, respectively.

Knockout *fabR* mutant strains were created by using the Lambda Red recombination method.[21] The *fabR* knockout allele was obtained from the *fabR* knockout mutant from the Keio collection, which has a kanamycin resistance (KanR) cassette replacing the gene.[23] A lambda Red recombination vector pKD78 was transformed into the *E. coli* B REL606 ancestor strain and selected on LB-chloramphenicol plates. These pKD78-containing strains were then induced with L-arabinose at 30°C and transformed with 100ng linear PCR product with the amplified KanR cassette from the *fabR* Keio strain (primers 133 and 134 in Table 1). The strains were then cured of the plasmid by growing at 37°C and cells with the knockout allele integrated into the chromosome were then selected for with kanamycin. Once the knock out allele was in the chromosome, the kanamycin cassette was subsequently cured from the genome to leave behind an in-frame deletion as previously described using the pCP20 vector.[21] The *fabR* knockout mutant was designated LNW25.

COMPETITIVE FITNESS ASSAYS

The method used for comparing fitness between two *E. coli* clones through head-to-head competition in co-culture has been previously described.[7,9,11,14] Each strain can be

differentiated by the Ara genetic marker (arabinose utilization) by plating on tetrazolium arabinose (TA) agar resulting in either red (Ara–) or white (Ara+) colonies. With one strain as Ara+ and the other as Ara– we combine them in the same culture, allow them to compete for resources, and then determine their relative fitness by plating and counting colonies of each color to determine the relative numbers of cells of each type.

For each competition assay, the strains to be competed were initially inoculated in LB from frozen stocks and allowed to grow up overnight with orbital shaking at 37°C before being transferred into DM25 via a 1:10,000 dilution for an additional overnight growth pre-conditioning in LTEE conditions. The following day, each strain was mixed at a 1:1 ratio with an overall 1:100 dilution and allowed to grow again for 24 hours in LTEE conditions. A dilution of the day 0 culture was plated before growth. On day 1, dilutions of these mixtures were again plated. Colonies were counted from both day 0 and day 1 plates. Realized growth rates of both competitors are determined using Malthusian parameters (m). If N_f is the final number of cells at the end of the experiment and N_i is the number of cells at the beginning of the experiment, then $m = \ln (N_f / N_i)$, is the Malthusian fitness of each strain. The relative fitness of two strains is the ratio of their realized Malthusian fitnesses during the competition experiment.

RNA-SEQ

Cultures of the ancestor, EW1, and EL1 strains were grown to exponential phase in DM25. In each case, total RNA was extracted (Promega SV total RNA isolation kit) and prepped as a strand-specific Illumina library (NEBNext Small RNA Library Prep).

Samples were sequenced at the UT Austin GSAF on a HiSeq instrument to obtain 100-base paired-end reads.

Data analysis of resulting FASTQ files was performed with the following steps: adaptor trimming, read alignment to the reference genome, and then read counting. Adaptors were trimmed using FLEXBAR,[24] reads were mapped using bowtie2,[25] counts of reads in genes were computed using a utility function of the *breseq* computational pipeline.[26] Resulting counts of reads matching each gene were then analyzed in R using DESeq[27] to compare expression between the ancestor strain and each evolved EW and EL strain.

RT-QPCR

To determine expression levels for *fabB* and *pspA* in the *fabR*30 mutant as compared to the ancestor, we grew cells under LTEE conditions for 24 hours. RNA was extracted (Promega SV total RNA kit) and reverse transcribed (Invitrogen SuperScript First-Strand Synthesis System for RT-PCR). To measure differences in *fabB* and *pspA* expression, we used *ihfB* as an internal control.[28] 5ng cDNA from the reverse transcription step was added to 2X Power SYBR® Green Master Mix (Applied Biosystems) and primers targeted to these three genes (*fabB*:137,138; *pspA*:139,140; *ihfB*:141,142), and run under standard PCR conditions on a ViiA™ 7 thermocyler (Life Technologies). Data were analyzed to determine ΔΔCT as previously described.[29]

FATTY ACID ANALYSIS

Fatty acid methyl esters (FAMEs) were purified using Bligh and Dyer lipid extraction from both the ancestor and *fabR* mutant strain LNW24.[30] These phospholipid samples were then sent to the Complex Carbohydrate Research Center (CCRC) at the University of Georgia for Gas Chromatography Mass Spectrometry (GC-MS) for lipid profiling. Acid methylation to isolate FAMEs was performed at CCRC prior to GC-MS.[31]

FLOW CYTOMETRY

Cells from both the ancestor and *fabR*30 strains were grown over 24 hours in DM100 and then diluted 1:10 in water. Diluted cells were then stained with 1 μg/ml FilmTracer™ 1-43 Green Biofilm Cell Stain (Life Technologies). In addition, cell number was determined using 135nM SYTO® 17 Red Fluorescent Nucleic Acid Stain (Molecular Probes) as a control as it stains all cells relatively equally. Samples were then run through a Fortessa flow cytometer and relative fluorescence and forward scatter patterns were compared.

Table 1. Primers used in this study

Primer #	**Sequence**	**Start site**	**End site**
1	TGC CGG ATG ATG GAA ACC TC	3893360	3893379
19	GAT GGC CTT CTT CAT GCA GG	3902369	3902350
20	AAC CAG TTT CAG ATC AAC CGG	3898963	3898943
P1 (89)	CAG TCA TAG CCG AAT AGC CT	Datsenko & Wanner[21]	
P2 (90)	CGG TGC CCT GAA TGA ACT GC	Datsenko & Wanner[21]	
k1 (93)	CAG TCA TAG CCG AAT AGC CT	Datsenko & Wanner[21]	
k2 (94)	CGG TGC CCT GAA TGA ACT GC	Datsenko & Wanner[21]	
kt (95)	CGG CCA CAG TCG ATG AAT CC	Datsenko & Wanner[21]	
108	TAA GCG GCA GGT TTC CGC TGT A	4140826	4140847
109	AGT AAA CCG TCA GCA CCA GGG A	4141776	4141755
117	CGC GGA TCC GTG CAA CCG CCG CCA ACA TT	4140515	4140534
118	CGC GGA TCC TCG GCA CCC GCA CTG AAG AC	4141486	4141565
133	AAG CGG CAG GTT TCC GCT GTA CGT AA	4140827	4140852
134	TGG AGC TAA ACA GCG CTG CGA AAG TA	4141725	4141700
135	CGG GTG GTA ACT TCG ACA AA	1033425	1033406
136	GCA TAA CCG GAT CGC CAA TA	1033327	1033346
137	GGG TGC GCT GTC TAC TAA ATA C	2384586	2384565
138	GCT CTT CAA CCA CTA CCA TAC C	2384471	2384492
139	CTG ATG ATC CAG GAG ATG GAA G	1366537	1366558
140	TGT TCA ATA CGG CGA GTA AG	1366634	1366618
141	AGA CGG TTG AAG ATG CAG TAA A	980978	980999
142	GCA AAG AGA AAC TGC CGA AAC	981079	981059
147	AAG CGG CAG GTT TCC GCT GTA	4140827	4140847
148	TGG AGC TAA ACA GCG CTG CGA AA	4141725	4141703

III. RESULTS AND DISCUSSION

IDENTIFY ADDITIONAL INSTANCES OF EVENTUAL WINNERS AND EVENTUAL LOSERS IN *E. COLI* POPULATIONS

IDENTIFY INDIVIDUAL LINEAGES BY RIBOSE OPERON PROFILING

In order to identify individual lineages with different beneficial mutations within evolving microbial populations, fitness-neutral or balanced genetic markers — such as fluorescent proteins[32] or the Ara marker[16,33] — are often utilized in order to differentiate between diverged subpopulations. Deletions that overlap the ribose (*rbs*) operon appear very early and often the LTEE populations due to the activity of a flanking IS*150* transposable element.[34] These deletions are slightly beneficial (~1% increased fitness) but much less so than other beneficial mutations that are driving adaptation at this point in the experiment (5-10% increased fitness).[5] Thus, *rbs* deletions can be thought of as a highly variable and fortuitous nearly neutral marker in the LTEE.[34]

There is heterogeneity in the sizes of ribose operon deletions between LTEE populations as well as within them (Fig. 4). A variety of ribose operon deletions are present in all 12 populations. These deletions occur when an IS150 element jumps into the operon and then recombines with the existing insertion sequence upstream of the ribose operon. This process could occur multiple times in one lineage, resulting in larger deletions as it evolves. So not every new deletion may represent a different lineage. Because multiple ribose deletions exist within a single population, genetically diverged

subpopulations can be identified by profiling ribose operon deletions. As they are nearly neutral, they serve as sentinels for detecting clades with secondary beneficial mutations.

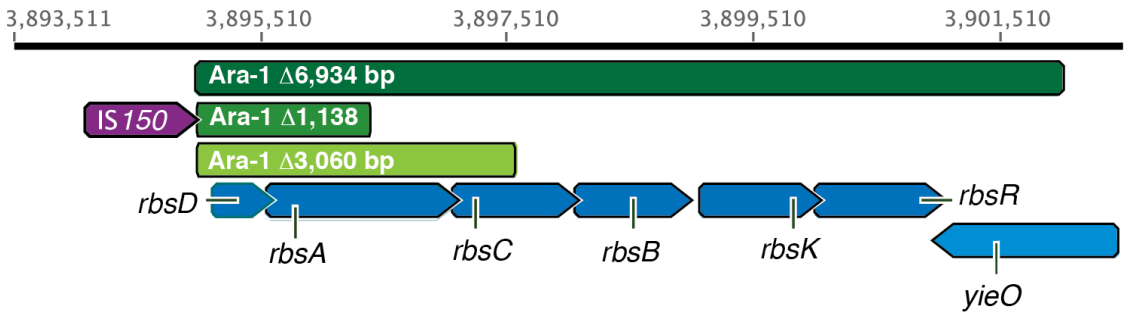

Figure 4. Various ribose operon deletions in LTEE populations.

These mutations result when a new IS element jumps into the ribose operon and recombines with the existing upstream IS element, resulting in deletions of various sizes, all with the same starting point. Evolved *E. coli* gain large deletions in the ribose operon (made up of 6 genes). The larger deletions include portions of downstream *yieO*, a putative drug efflux gene.[34]

Previous work isolated clones in Ara–1 at the earliest LTEE time points and identified three subpopulations using the size of the deletion in the ribose operon to distinguish them (Fig. 5a). The eventual winners and losers coexist at 500 generations and the eventual winners (Δ6934 bp) fix by 1500 generations. To identify subpopulations in two additional populations, Ara–5 and Ara–3, I profiled ribose operon deletion heterogeneity. Within the Ara-5 population only four deletions were found (Fig. 5b). From these one (Δ5627 bp) was prominent early in evolution at 500 generations and eventually fixed in the population by 1500 generations.

FIGURE 5 (needs formatting)

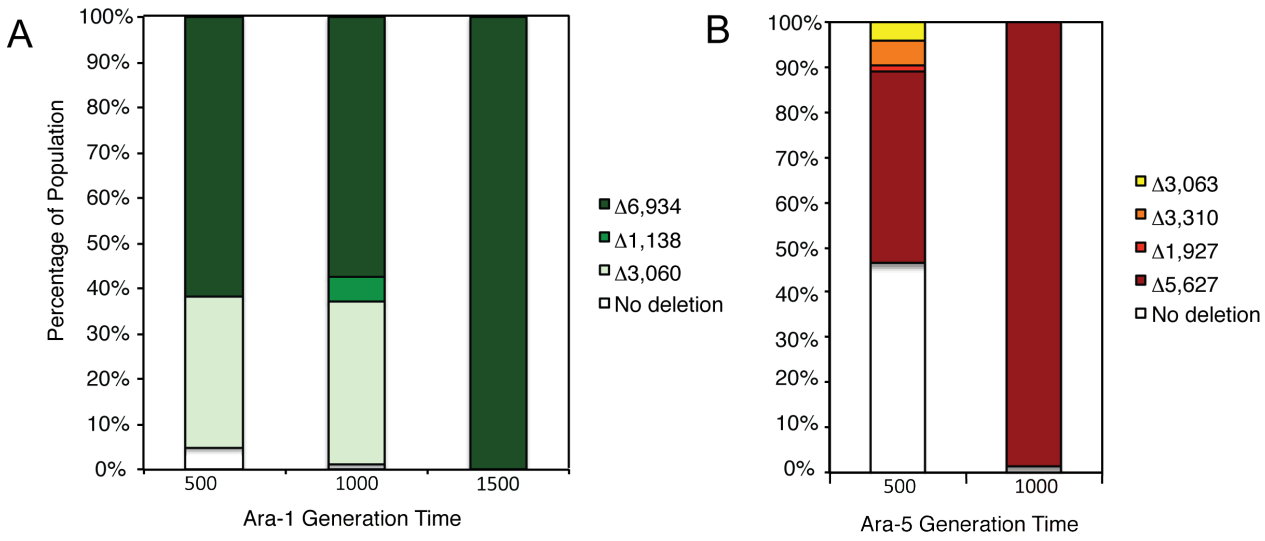

Figure 5. Heterogeneity of ribose deletions within populations.

As in Figure 4, there are a variety of ribose deletions with different sizes within populations. Each color represents a different deletion. It is likely that each deletion originated as an independent deletion event in the context of the ancestral ribose operon sequence but possible that an existing deletion may be enlarged by another round of the same IS element recombination process at later generations. (A) Analysis of population Ara–1 revealed only three ribose deletion sizes, one of which (Δ6934 bp) was already prominent at 500 generations and fixed by 1500 generations. (B) Analysis of population Ara–5 found four deletion sizes, and the deletion that almost fixed by 1000 generations (Δ5627 bp) was already prominent within the population by 500 generations.

However, Ara-3 showed a highly heterogeneous mix of early ribose deletions (Fig. 6a). One lineage identified in this way (Δ6406 bp) seems to come very close to fixing within the population: at 1000 generations none of the 44 clones tested were found with any other deletion. However, the Ara–3 clade that eventually won has a larger ribose deletion. This cannot be a case where the deletion size increased due to a second IS-mediated deletion event because these isolates lack a known mutation in *fabR* (discussed below). Furthermore, we expect these types of secondary deletion events to be much rarer

because there is likely only a fitness advantage for knocking out *rbs* operon function for the first such deletion.[34] The phylogenetic relationships between the various *rbs* deletions and the *fabR* mutation are represented in figure 6b. Ara-3 seems to be a promising example of a population where one lineage is most fit for over 2000 generations, but its trajectory was a dead end and resulted in eventual extinction.

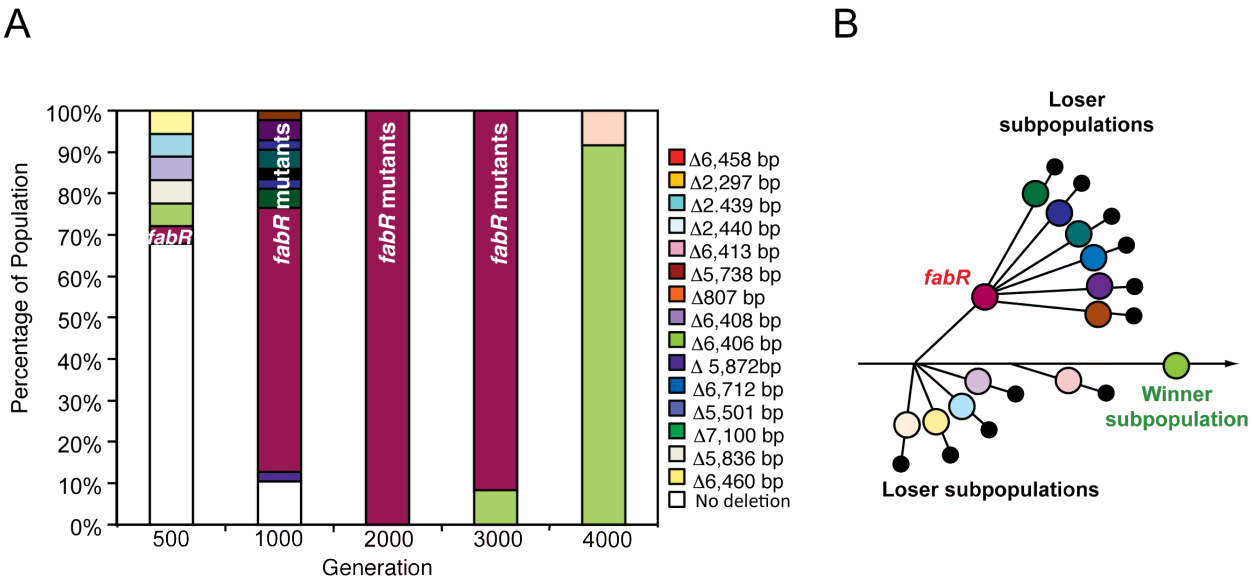

Figure 6. Heterogeneity of ribose deletions within population Ara–3.

(A) Fifteen different deletions were identified in this population. The one that wins (Δ6406 bp) in lime green existed at 500 generations but then drops to an unperceivable level until 3000 generations. The 6406-bp deletion is present in 100% of the population by 5000 generations. During the early generations a different deletion (Δ5738 bp) in maroon seems to almost fix, but surprisingly it goes extinct instead. This example and other unsuccessful ribose deletions occurred in the same genetic background with a *fabR* allele (as labeled). (B) The relationship between various ribose deletions and an evolved *fabR* mutation in Ara–3 is depicted in a phylogenetic tree.

MUTATIONS IN *FABR* ARE COMMON IN EVENTUAL LOSERS IN OTHER POPULATIONS

To understand patterns in evolution across all 12 parallel populations we sequenced mixed population samples at the earliest time points, looking for genes that are repeatedly mutated early in the LTEE. The whole-genome resequencing data from populations at 500 and 1000 generations revealed a number of candidate genes that may be responsible for fitness increases in some of these lineages. Although a number of mutations were revealed through these data, the only gene with mutations in most of the populations at these early time points was *fabR*. This was of particular interest since there were not only mutations in *fabR* in 7 out 12 populations early in their evolution, but a *fabR* mutation had already been found in the previously studied eventual loser population in Ara-1 (Fig. 7).[16] As in Ara-1, mutations in *fabR* occur early and go extinct in the other populations. The only populations in which a mutation in *fabR* exists late in evolution (40,000 generations) had become hypermutators by that point. Thus, the new mutations in *fabR* at these time points likely don't provide the same sort of fitness benefit. Instead they may result from genetic drift due to the greatly elevated genome-wide mutation rate. Mutations occurring in *fabR* include 8 nonsynonymous point mutations, one nonsense mutation, and a single base pair insertion, (Fig. 7). Since two of these events result in a mid-gene stop codon or frameshift, these data suggest that loss of FabR function is beneficial in the environment of the LTEE.

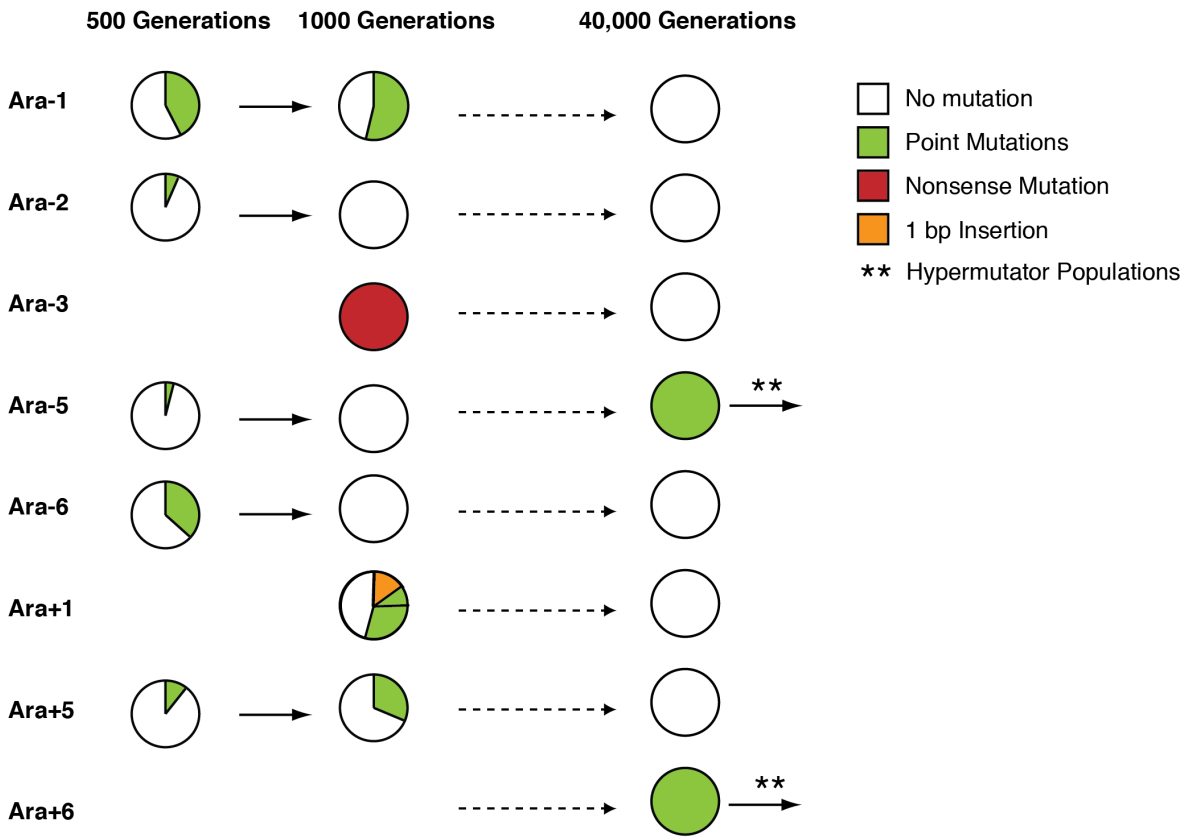

Figure 7. Mutations in *fabR* are prevalent in LTEE populations.

Metagenomic sequencing of early time points for all 12 populations in the evolution experiment revealed various mutations present at 500 and 1000 generations. The gene that had mutations in the highest percentage of populations at these time points was *fabR;* 7 of the 12 populations had observable lineages with a mutation in this gene. Interestingly, these *fabR* lineages always go extinct. Asterisks denote populations that evolved to be hypermutators by this point in the experiment. Unlike for the earlier alleles, the presence of *fabR* mutations in these samples is likely a result of a heightened mutation rate rather than a beneficial effect on fitness.

ELUCIDATE THE PHYSIOLOGICAL EFFECTS OF BENEFICIAL *FABR* MUTATIONS

To verify that mutations in *fabR* are beneficial early in the LTEE, I constructed a strain with the mutant *fabR* allele (T30N) in an ancestral strain background. In addition, due to the expectation that *fabR* mutations result in a loss of function in the FabR protein,

I also constructed a *fabR* deletion strain to test for comparison. Each *fabR* mutation resulted in a ~4% increase in fitness in co-culture competition assays versus the ancestral strain (Fig. 8). The Ara-1 population's *fabR* mutant allele had fitness effects similar to the knockout mutant. This result suggests that each early *fabR* mutation that we observe in the evolving strains is beneficial and that they result in a loss of FabR repressor function.

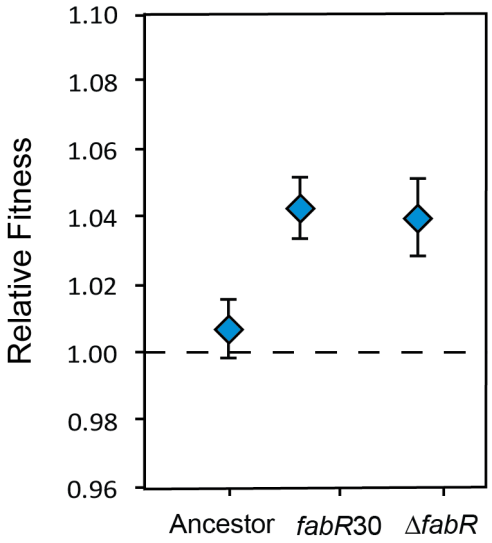

Figure 8. Fitness benefits of mutations in *fabR*.

Both the Ara-1 *fabR* mutant allele and the *fabR* knock out mutant have increased relative fitness as compared to the ancestor when competed in LTEE conditions. Error bars are 95% confidence limits.

The fitness benefit conferred by *fabR* mutations somehow results from blocking a functional FabR unsaturated fatty acid synthesis repressor. Under normal conditions, as the cell requires unsaturated fatty acids (UFAs) a number of genes encoding synthesis enzymes are expressed (Fig. 9a) However, as mentioned above, FabR responds to growth conditions where additional unsaturated fatty acids are not required by dimerizing and

forming a complex with ACP covalently attached to long UFAs. To repress the synthesis machinery, FabR binds to the promoter regions of two genes required for UFA synthesis, *fabA* and *fabB*, turning off UFA production (Fig. 9b).

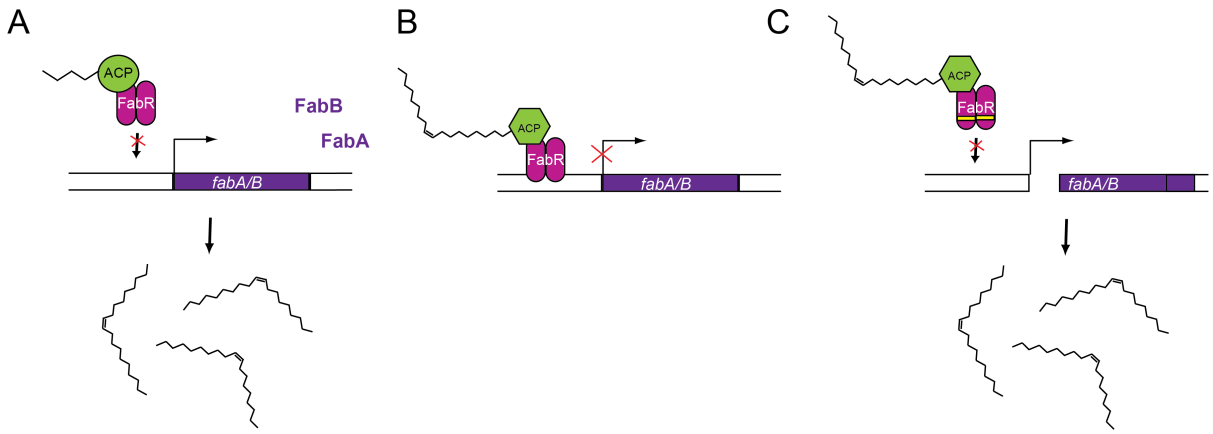

Figure 9. FabR maintains lipid homeostasis by repressing unsaturated fatty acid synthesis.

(A) In actively growing cells, acyl carrier protein (ACP) tends to have a higher ratio of saturated chain molecules attached to it because unsaturated chains are continually being used as a substrate for the synthesis of new lipid molecules. When this form of ACP is present, it sequesters the FabR repressor and allows expression of *fabA* and *fabB*. (B) In starving cells, ACP attached to unsaturated chain molecules accumulates. This form of ACP does not inhibit FabR, so it is able to bind to the promoters for *fabA* and *fabB*, genes and block their expression. (C) Loss-of-function mutations in FabR presumably lead to *fabA* and *fabB* transcription regardless of nutrient conditions.

A nonfunctional FabR protein would likely no longer retain the ability to respond to UFA-sufficient conditions and would no longer appropriately repress the transcription of *fabA* and *fabB* (Fig. 9c), presumably resulting in an increased ratio of UFAs in these mutants. Some of the expected downstream effects of a non-functional FabR repressor can be shown in these mutants. There is a clear increase in expression levels of FabB in the *fabR* T30N mutant as compared to the ancestor during exponential growth (Fig. 10).

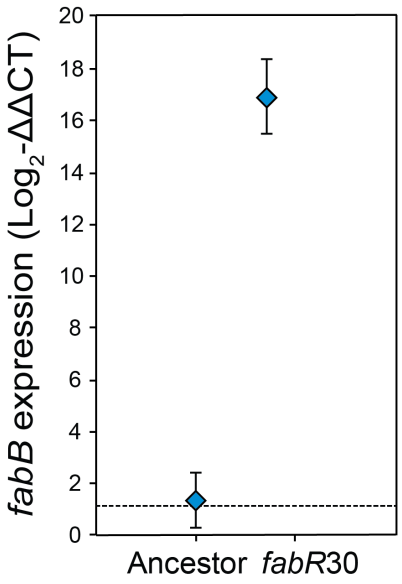

Figure 10. Mutations in *fabR* upregulate fatty acid synthesis enzyme *fabB*.

The ancestor and fabR30 strains were grown in LTEE conditions to mid-log phase, harvested. After RNA was isolated and reverse transcribed, RT-qPCR for *fabB* was performed, using *ihfB* as an internal control. Results show a significantly higher level of expression in the *fabR* mutant as compared to the ancestor, suggesting that mutations in *fabR* result in drastic up-regulation of genes under its control.

To test whether the lipid composition is altered due to a mutant *fabR* allele, fatty acid methyl esters (FAMEs) were profiled by GC-MS or the ancestral strain as well as the *fabR* T30N mutant strain (Fig. 11). Notably, the unsaturated fatty acid chain of oleic acid (18:1) increases from 12.0% to 22.2% relative molar abundance from the ancestor to the *fabR* mutant strain, as expected. These findings are consistent with changes in lipid composition observed in *fabR* mutants of other *E. coli* strains.[18]

Glycosyl Residue	Ancestor (% mol)	*fabR* (% mol)
Myristic acid (14:0)	6.9	4.8
Palmitoleic acid (16:1)	58.1	55.3
Palmitic acid (16:0)	13.7	10.1
Oleic acid (18:1)	12.0	22.2
Heptadecanoic acid (17:0)	8.9	7.3
Stearic acid (18:0)	0.3	0.4
Total	100	100

Figure 11. Lipid composition in the *fabR* mutant versus the ancestor.

The ancestor and fabR30 strain were grown up to mid-log phase in DM1000. Total phospholipids were isolated using the Bligh-Dyer lipid isolation protocol and then acid methylated to produce fatty acid methyl esters appropriate for GC-MS analysis.[30,31] Unsaturated fatty acid content, particularly oleic acid, increased in *fabR* mutants, corresponding to previous research.

UNDERSTAND WHY EVENTUAL LOSERS ONLY EXIST EARLY IN EVOLUTION IN THE LTEE

WHY ARE *FABR* MUTATIONS BENEFICIAL IN THE LTEE?

Given the evidence that mutations in *fabR* result in a loss of repressor function causing continued enzyme synthesis and, thus, increased levels of UFAs during conditions when the synthesis machinery may have otherwise been turned off, we would like to determine why this phenotype is beneficial under the conditions of the LTEE. One idea is that more lipids could enhance the available building blocks for the cell membrane and increase cell size. Previous research on the LTEE populations has shown that the

cells do indeed increase in size as they evolve.[7] While a larger cell size could have resulted from the more unsaturated FAs generated in a *fabR* mutant strain, our results suggest that this is not the case. Figure 12 indicates a clear change in cell size when comparing the ancestral strain to one isolated after 40,000 generations. However, the *fabR* mutant does not seem to increase in size in a significant manner. This suggests that increased FA synthesis or the related downstream regulatory effects of a *fabR* mutation are benefitting these strains in some other way.

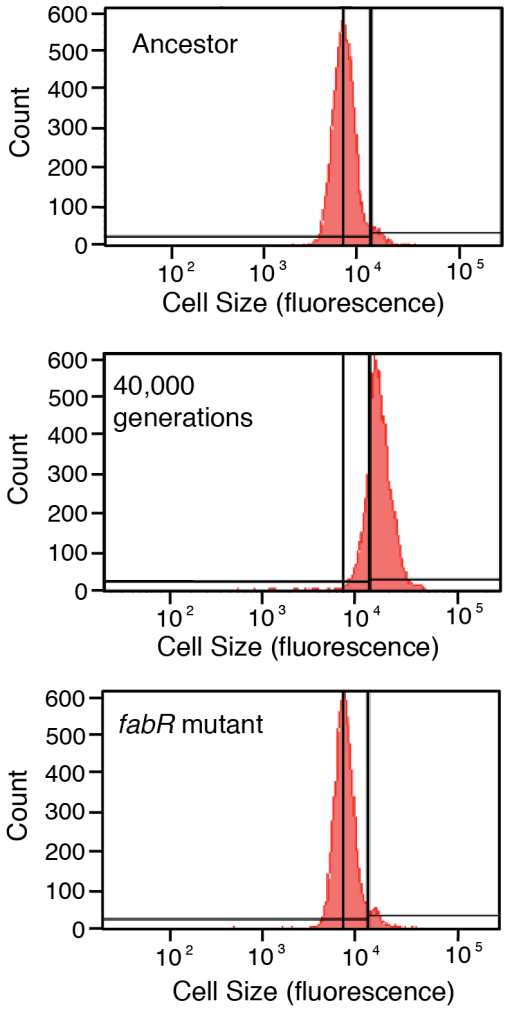

Figure 12. Mutations in *fabR* do not affect cell size.

After overnight growth in LTEE conditions, *E. coli* cells with their plasma membranes stained with a membrane-specific fluorescent dye were run through a flow cytometer to measure fluorescence per cell. Larger cells, such as those at from generation 40,000 of the LTEE, used here as a positive control will have most of the cell counts at a higher fluorescent measurement.[7] However, *fabR* mutant strains have a pattern nearly identical to ancestors and therefore seem to be the same size.

Another possibility is that increased fatty acid synthesis in a minimal glucose environment such as the LTEE could turn off or delay specific starvation-induced processes that halt cellular replication. The stringent response is the best-studied of these responses in *E. coli* and has been shown to halt cell growth in minimal glucose environments.[20] There are several pathways associated with the stringent response that eventually lead to a global regulatory protein, either RelA or SpoT, synthesizing the alarmone (p)ppGpp, which stops ribosome biogenesis and, thus, cellular replication. If altering the sensitivity of the stringent response is involved in the evolution of LTEE lines it is most likely affecting the SpoT pathway, which responds to both glucose starvation as well as fatty acid starvation.[20] Since FabR repression is turned off in *fabR* mutants, as demonstrated by RT-qPCR of FabB above, then the increase in fatty acids could alter the effect of glucose starvation in these lineages. Thus, we might expect that cells would slow the activation or action of the stringent response and grow faster, continue to grow, or decrease lag time. All of these possibilities could result in a higher fitness as compared to the ancestor strain.

Accordingly, Figure 13a shows the increase in growth rate during exponential phase in *fabR* mutants. Somewhat less clearly, we see an effect on the lag time in mutants (Fig. 13b), but these results are less convincing. Regardless, there is a beneficial effect on growth dynamics in the *fabR* mutants. One way to directly test the involvement of the stringent response would be to quantify the relative levels of (p)ppGpp in the mutants as compared to the ancestor. Presumably, if the stringent response is not as robust in mutant strains we would see a decrease in (p)ppGpp levels. However, my results for this test

were inconclusive (data not shown). In addition, previous studies of the LTEE and the effects of *spoT* mutations were not able to detect changes in (p)ppGpp levels.[14]

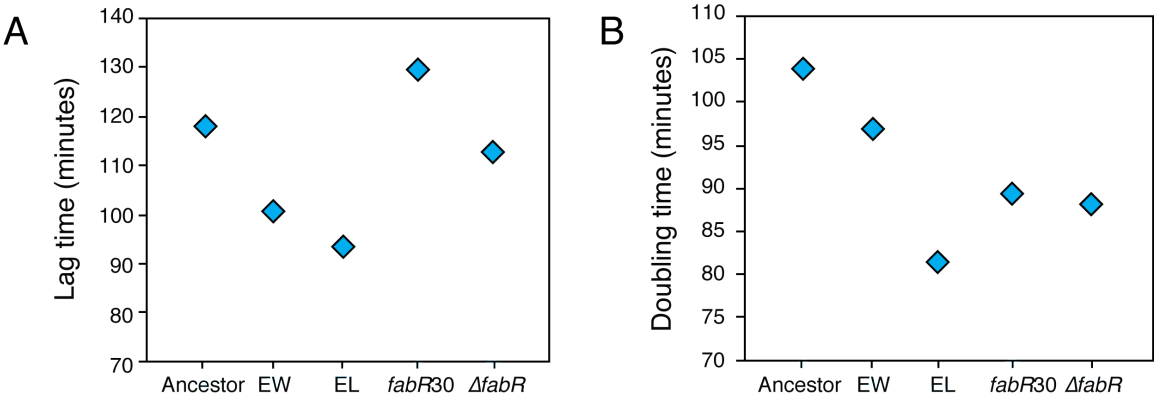

Figure 13. Evolved strains have altered growth dynamics.

Mutations in *fabR* appear to contribute to shorter doubling times in the eventual losers. Strains were grown under LTEE conditions for 24 hours. The following day, cells were transferred to a 96 well plate in triplicate at a 1:100 dilution in DM25. Spectrophotometry readings at 420 nm were recorded for all wells every 10 minutes for 6 hours in a microplate reader. These growth curves were then fit to a standard growth model in R to estimate the lag phase duration and doubling time. As expected, the evolved clones (EW and EL from the 500 generation time point in Ara-1) had shorter lag phases and doubling times. However, constructed *fabR* mutants seemed only to improve their doubling times, suggesting that this mutation (which is present in EL, but not EW) accounts for some of the improvement in EL.

In preliminary RNA-Seq results used to analyze gene expression levels in the eventual losers and eventual winners at 500 generations, I found evidence supporting the increased *fabA* and *fabB* expression levels in the eventual loser, which contains a *fabR* mutation (Fig. 14). In addition to these somewhat expected up-regulated genes in lineages with *fabR* mutations, we also found increased expression in a number of other genes. Of particular interest is increased expression of the membrane damage response regulator, *pspA* (Fig. 14). Because *fabR* mutations would alter unsaturated versus saturated FA

content in the cell, this could result in a damaged cell membrane. PspA is part of the membrane stabilizing system in *E. coli* and is normally induced in response to membrane stress.[35] These results were confirmed with RT-qPCR analyzing *pspA* expression in the *fabR* mutant as compared to the ancestor (Fig. 15).

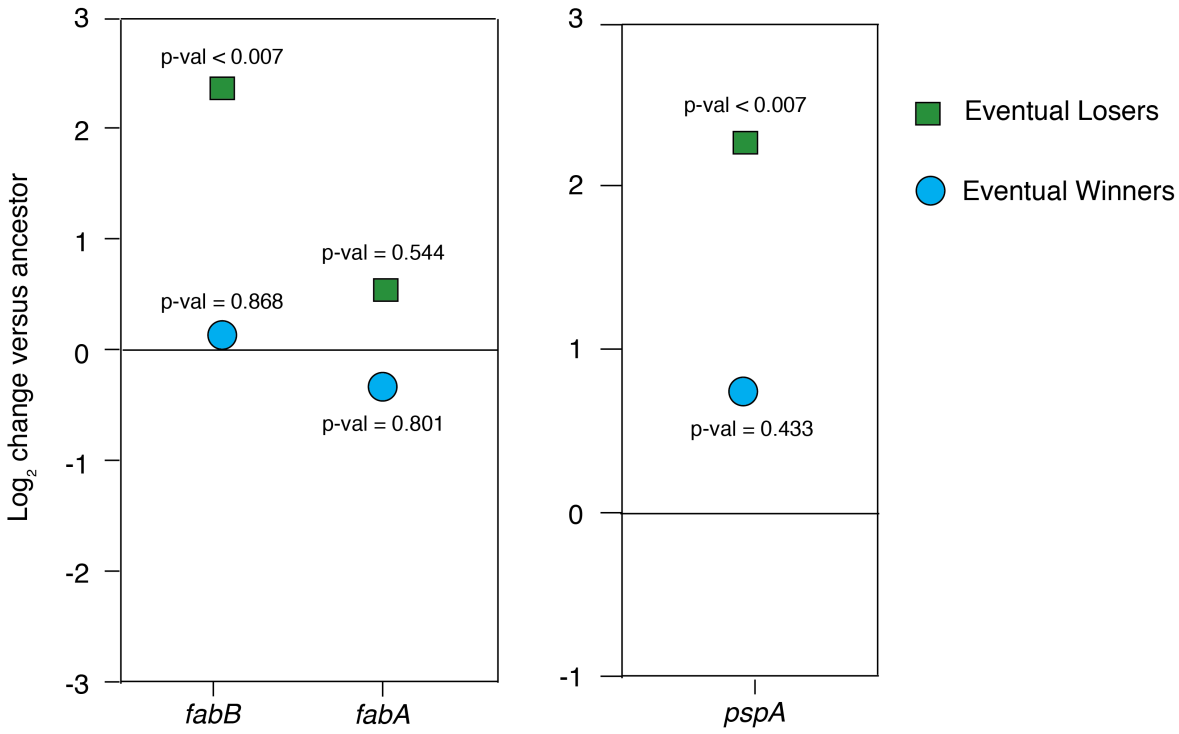

Figure 14. RNA-Seq in eventual losers and winners.

Eventual losers exhibit increased expression of both fatty acid synthesis genes as well as membrane damage response genes. Ara−1 eventual winner and loser clones and the ancestral strain were grown in LTEE conditions then transferred and grown to mid-log phase in DM25 before cells were harvested for RNA isolation. Illumina RNA-Seq confirmed that eventual losers with *fabR* mutations have increased expression of the transcripts that FabR represses (*fabA* and *fabB*). In addition, these clones have significantly increased mRNA levels for the membrane damage response gene, *pspA*, suggesting side effects on the membrane from altered fatty acid levels. *P*-values are for a significant difference versus the ancestor.

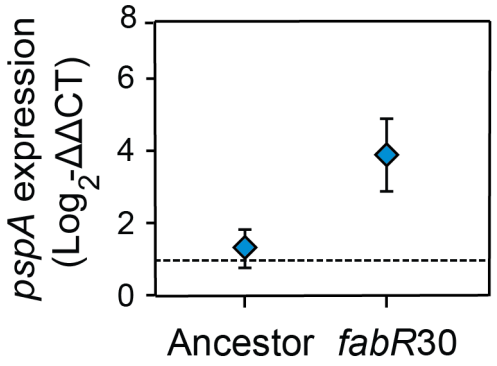

Figure 15. *fabR* mutations increase *pspA* expression in eventual losers.

The ancestor and *fabR*30 strains were grown in LTEE conditions to mid-log phase. After RNA was isolated and reverse transcribed, RT-qPCR for *pspA* was performed, using *ihfB* as an internal control. Results show a significantly higher level of expression in the *pspA* mutant as compared to the ancestor, suggesting that mutations in *fabR* result directly result in the increased expression of the membrane damage response gene previous observed in the eventual loser though RNA-Seq.

Due to the potentially destabilizing effects of an altered lipid ratio that may result in a damaged or more permeable membrane in these mutants, I wanted to see if they would be more prone to environmental stresses. Although *fabR* mutations are clearly beneficial early on and can expand to large proportions of the population, they eventually lead these lineages to extinction, so this could be an explanation for a disadvantage to acquiring a *fabR* mutation. However, when the ancestor and *fabR* mutant were subjected to increasing NaCl concentrations, the effects were essentially identical between the two strains (Fig. 16). Similar results were obtained when the ancestor and *fabR* mutants were grown at several different temperatures. In theory more unsaturated fatty acids will increase membrane permeability and this could be detrimental at higher temperatures but better for growth at lower temperatures. However, the *fabR*30 mutant exhibited roughly

the same fitness benefit versus the ancestral strain at 25°C and 30°C (Fig. 17). We were unable to get consistent results when we attempted to measure relative fitness at 42°C (large error bars).

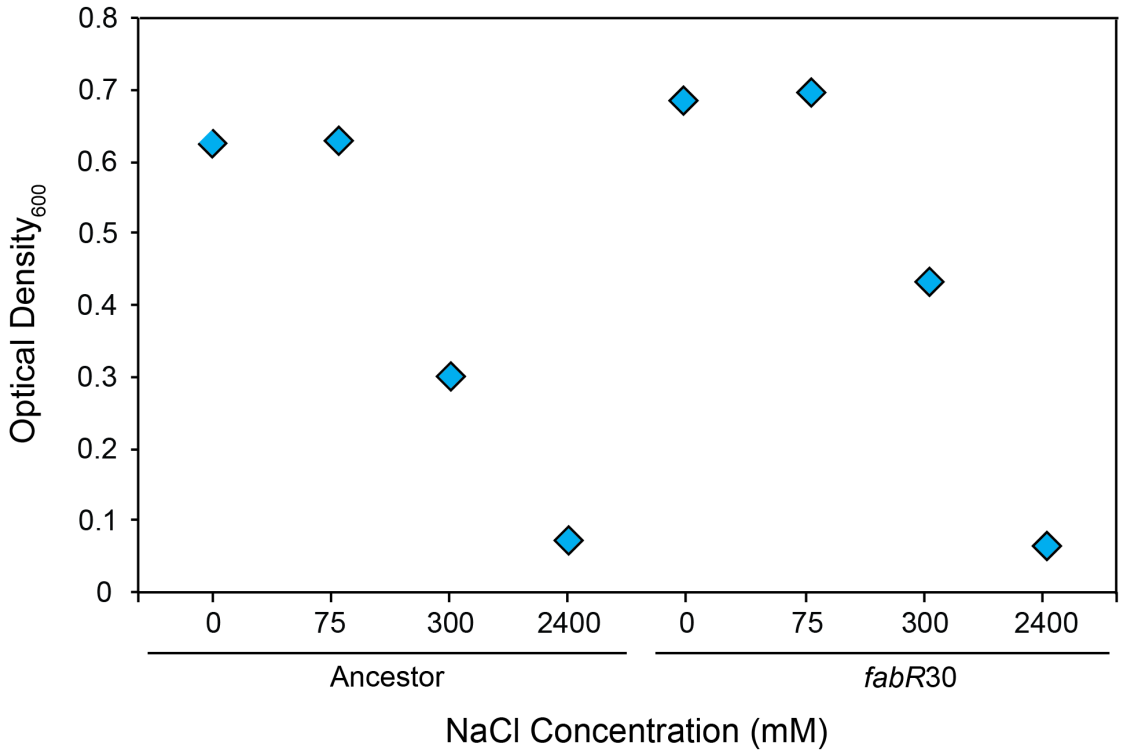

Figure 16. Environmental salt stress does not affect the growth of *fabR* mutants.

The optical densities at 600 nm of cultures after 24 hours of growth in DM25 with various NaCl concentrations were measured. Although mutant membranes may be affected by altered lipid compositions, this does not lead to susceptibility from some environmental stress, such as salt concentrations.

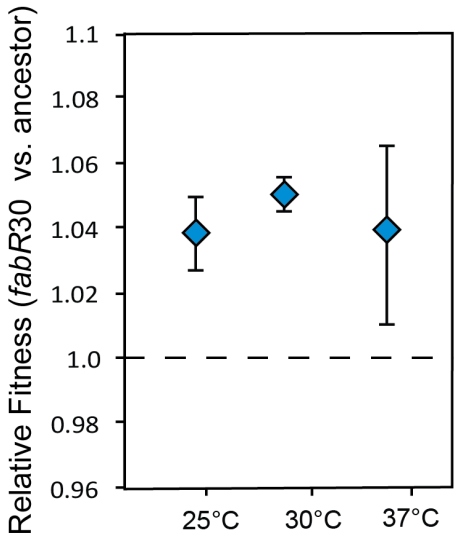

Figure 17. Temperature does not affect fitness of *fabR* mutant.

Strains were competed at a 1:1 ratio in different temperatures, and shown to have no significant change in relative fitness at various temperatures. Error bars are 95% confidence limits.

UNDERSTAND HOW *FABR* MUTATIONS INTERACT WITH OTHER POTENTIAL BENEFICIAL MUTATIONS

The prevalence of *fabR* mutations in LTEE populations in early evolutionary stages and the eventual extinction of all these lineages suggests that the benefits from *fabR* mutations are not enough for fixation of these lineages. There is a pattern across populations of certain early successful clades going extinct, possibly as a consequence of epistatic interactions between early mutations. If the loss of FabR function is beneficial because it delays the onset of the stringent response, its inevitable extinction is somewhat surprising. I would like to determine if this mutation interacts with other processes and whether alternative mutations in the eventual winners confer the same physiological benefits, and how they may achieve the same effects without restricting further evolution.

Despite the fitness benefits conferred by *fabR* mutations, these mutations go extinct in all populations. Therefore, there must be some interaction with other processes to block the ability of these lineages to fix within the populations. In several populations, including Ara-1, a mutation in *spoT* occurs later in the eventual winners. In fact, there are 15 known mutations that occurred in *spoT* in the LTEE or the EW/EL replay experiments (Fig. 18).

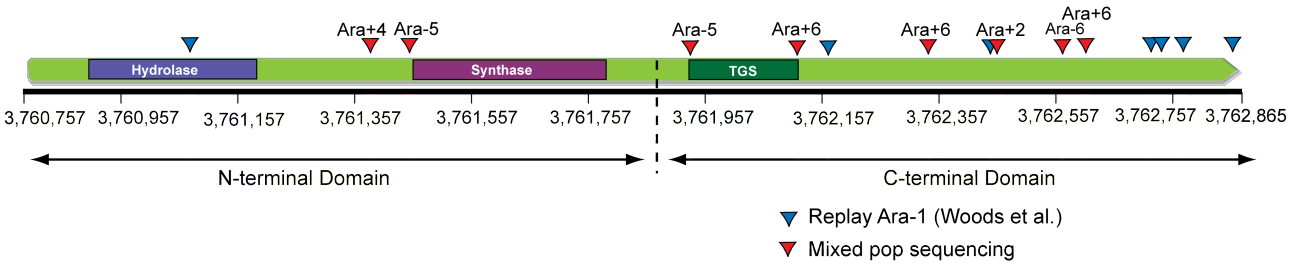

Figure 18. Mutations occur in *spoT* in most populations in the LTEE.

Various experiments have identified *spoT* mutations and these mutations are not isolated to a single protein domain. Blue triangles represent *spoT* mutations previously identified in replay experiments that began with EW and EL Ara–1 isolates [16] while red triangle were found in whole population resequencing data in 500 and 1000 generations in this study.

If *fabR* mutations affect the stringent response, we might also believe that *spoT* mutations do the same. In fact, since SpoT is a specific regulator for this cellular response it may be a more direct route to alter the same regulatory process. As mentioned above, during conditions of glucose or fatty acid starvation *E. coli* typically activate the stringent response. Acyl carrier protein (ACP) is covalently attached to short fatty acid chains under these conditions. This may indicate a decrease in full-length fatty acid synthesis due to lack of resources, which can be recognized by SpoT. This recognition activates the

synthesis activity of SpoT and (p)ppGpp is made, turning on the stringent response which halts ribosome biogenesis (Fig. 19a). When full-length fatty acids are abundant, ACP will be attached to them, causing a conformational change in ACP and causing it to no longer activate SpoT (Fig. 19b). Likewise, in *fabR* mutant strains that make more fatty acids even when their synthesis should be repressed due to a nutrient-limited environment, they will follow the path shown in Figure 19b and decrease or delay the stringent response. However, Figure 19c shows that in lineages with mutations in *spoT*, even though only short fatty acid intermediates are bound to ACP, the *spoT* mutation could block their ability to activate the (p)ppGpp synthesis activity.

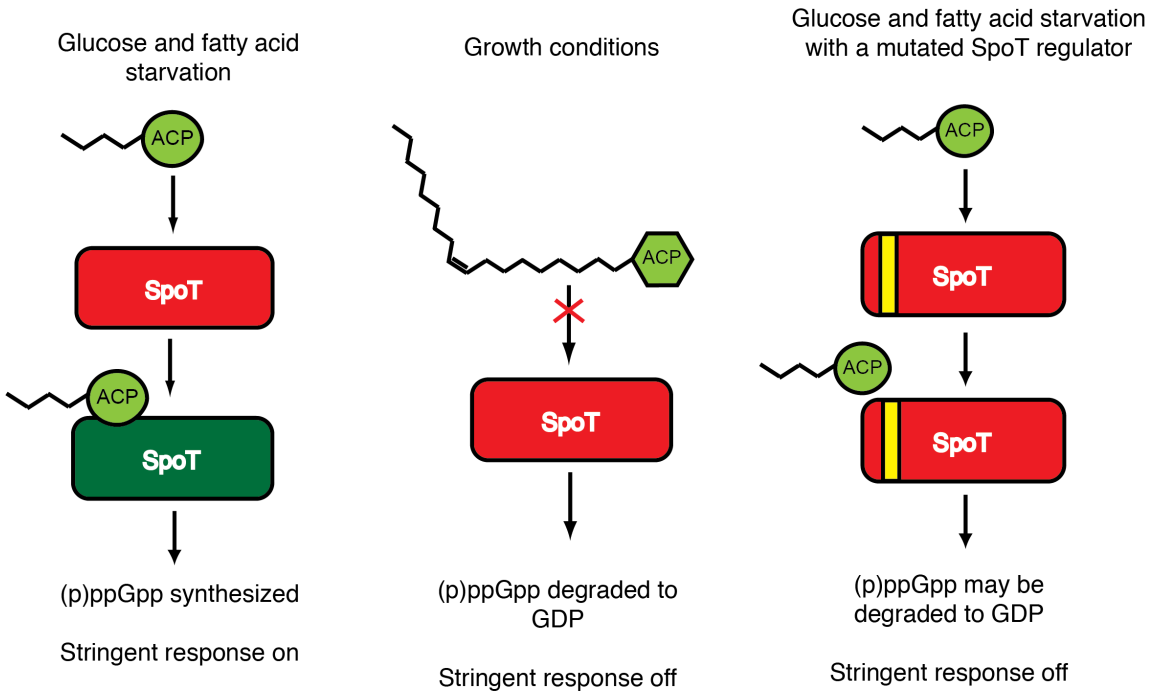

Figure 19. Mutations in *spoT* can affect normal stringent response to stressful environments.

(A) Normal stringent response and growth arrest are turned on during glucose and fatty acid starvation. (B) During times when glucose and fatty acids are plentiful SpoT degrades (p)ppGpp instead of synthesizing it and the cell grows normally. This may also occur if *fabR* mutations result in fatty acid synthesis during times of starvation. (C) The stringent response may also by avoided or delayed if mutations occur in *spoT*, resulting in altered SpoT function.

This model leads to several predictions. If both mutations negatively affect the ability of the cell to activate the stringent response through the same downstream pathway, I would expect cells with a *fabR* mutation would not gain a fitness benefit from an additional *spoT* mutation. It is possible that *spoT* mutations (or other mutations affecting the stringent response) reproducibly win over *fabR* mutations because they are direct hits on the master regulator of the stringent response as opposed to an upstream

mutation that has an equivalent effect on the stringent response but also affects other metabolic processes (fatty acid biosynthesis) in ways that may have negative side effects that restrict further evolutionary pathways.

To date, I have found no clones containing both *fabR* and *spoT* mutations. They appear to be mutually exclusive adaptive steps. I hypothesize that both *fabR* and *spoT* mutations delay the stringent response and that because SpoT is specific to the stringent response it is less likely to have detrimental side effects that could result from an alteration in a cell's lipid profile. Mutations in *fabR* should increase the unsaturated FAs in the cell with possible negative side effects on membranous structures that utilize FAs, while mutations in *spoT* should not have that affect while maintaining the potential benefit of delaying the stringent response.

To test the other effects of *fabR* and *spoT* mutations, first I measured fitness in strains containing *fabR*, *spoT*, and both mutations. If there are functional interactions between *fabR* and *spoT* mutations, a combination that does not seem to occur naturally, they should not have greater fitness than just one of the mutations alone. Indeed, fitness benefits of the double mutant were not greater than each single mutant alone and may have actually decreased in fitness (Fig. 20). This lack of cumulative fitness benefit in the double mutant suggests an overlap in cellular effects resulting from both mutations.

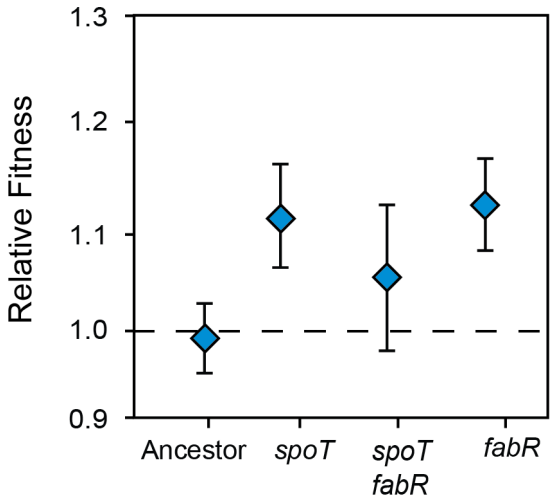

Figure 20. Mutations in *spoT* and *fabR* do not result in a cumulative fitness increase.

Both single mutants of fabR30 and the *spoT* allele that eventual fixes in Ara−1 as well as the double mutant were competed in co-culture against the ancestor. The results indicate that the double mutant is not more fit than the single mutants. This result might be expected if both mutations had downstream effects on the same cellular process that was important for fitness.

CONCLUSIONS

While the downstream cellular processes affected by *fabR* and *spoT* mutations are still not understood, I believe that they both affect the stringent response. Ideally, the most direct evidence for this would be a decreased ratio of (p)ppGpp:GTP levels. Unfortunately, attempts at these experiments never resulted in a definite answer. If is likely though, based on previous research, that even if these experiments had worked, they may not have resulted in a clear answer since *spoT* mutations in other LTEE populations don't seem to alter (p)ppGpp ratios.[14] If the stringent response is behind *fabR* and *spoT* mutational benefits in these lines, it may be that a subtle change is enough to have the beneficial effects on cellular growth. Alternatively, an uncharacterized regulatory activity of *spoT* may be key to mediating the observed effects, as the *spoT* mutations nearly all occur outside of the hydrolase and synthase domains in this protein (Fig. 19). Regardless of the exact cellular mechanism, the consistent extinction of those lineages that acquire mutations in *fabR*, provides evidence that the eventual winner and eventual loser evolutionary dynamics are a fairly common phenomenon in evolving large asexual populations such as bacteria. Hopefully, further developments that build on this work will lead toward a better understanding of the intricacies of natural selection in this system and the importance of adaptive dead ends in fitness landscapes in general.

REFERENCES

1. Chao, L. & Cox, E. C. Competition between high and low mutating strains of Escherichia coli. *Evolution* **37,** 125–134 (1983).

2. Arjan G., J. Diminishing Returns from Mutation Supply Rate in Asexual Populations. *Science* **283,** 404–406 (1999).

3. Lang, G. I., Botstein, D. & Desai, M. M. Genetic variation and the fate of beneficial mutations in asexual populations. *Genetics* **188,** 647–661 (2011).

4. Lang, G. I. *et al*. Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature* **500,** 571–574 (2013).

5. Khan, A. I., Dinh, D. M., Schneider, D., Lenski, R. E. & Cooper, T. F. Negative epistasis between beneficial mutations in an evolving bacterial population. *Science* **332,** 1193–1196 (2011).

6. De Visser, J. A. G. M., Cooper, T. F. & Elena, S. F. The causes of epistasis. *Proc. R. Soc. B Biol. Sci.* **278,** 3617-3624 (2011).

7. Lenski, R. E. Phenotypic and genomic evolution during a 20,000 generation experiment with the bacterium Escherichia coli. *Plant Breeding Reviews* **24,** 225-265 (2004).

8. Papadopoulos, D. *et al*. Genomic evolution during a 10 , 000-generation experiment. *PNAS* **96,** 3807–3812 (1999).

9. Lenski, R. E., Rose, M., Simpson, S. & Tadler, S. Long -term experimental evolution in Escherichia coli. I. Adaptation and divergence dueing 2,000 generations. *Am. Nat.* **138,** 1315–1341 (1991).

10. Cooper, T. F., Rozen, D. E. & Lenski, R. E. Parallel changes in gene expression after 20,000 generations of evolution in Escherichiacoli. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 1072–1077 (2003).

11. Lenski, R. E. & Travisano, M. Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. *Proc. Natl. Acad. Sci. U. S. A.* **91,** 6808–14 (1994).

12. Wiser, M. J., Ribeck, N. & Lenski, R. E. Long-term dynamics of adaptation in asexual populations. *Science* **342,** 1364–7 (2013).

13. Barrick, J. E. *et al*. Genome evolution and adaptation in a long-term experiment with Escherichia coli. *Nature* **461,** 1243–7 (2009).

14. Philippe, N., Crozat, E., Lenski, R. E. & Schneider, D. Evolution of global regulatory networks during a long-term experiment with Escherichia coli. *Bioessays* **29,** 846–860 (2007).

15. Blount, Z. D., Borland, C. Z. & Lenski, R. E. Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 7899–906 (2008).

16. Woods, R. J. *et al*. Second-order selection for evolvability in a large Escherichia coli population. *Science* **331,** 1433–6 (2011).

17. Fujita, Y., Matsuoka, H. & Hirooka, K. Regulation of fatty acid metabolism in bacteria. *Mol. Microbiol.* **66,** 829–39 (2007).

18. Zhu, K., Zhang, Y.-M. & Rock, C. O. Transcriptional regulation of membrane lipid homeostasis in Escherichia coli. *J. Biol. Chem.* **284,** 34880–8 (2009).

19. Angelini, S., My, L. & Bouveret, E. Disrupting the Acyl Carrier Protein/SpoT interaction in vivo: identification of ACP residues involved in the interaction and consequence on growth. *PLoS One* **7,** e36111 (2012).

20. Battesti, A. & Bouveret, E. Acyl carrier protein/SpoT interaction, the switch linking SpoT-dependent stress response to fatty acid metabolism. *Mol. Microbiol.* **62,** 1048–63 (2006).

21. Datsenko, K. a & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 6640–5 (2000).

22. Cooper, T. F., Rozen, D. E. & Lenski, R. E. Parallel changes in gene expression after 20,000 generations of evolution in Escherichiacoli. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 1072–1077 (2003).

23. Baba, T. *et al*. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2,** 2006.0008 (2006).

24. Dodt, M., Roehr, J., Ahmed, R. & Dieterich, C. FLEXBAR—Flexible Barcode and Adapter Processing for Next-Generation Sequencing Platforms. *Biology (Basel).* **1,** 895–905 (2012).

25. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9,** 357–9 (2012).

26. Barrick, J. E. & Lenski, R. E. Genome-wide mutational diversity in an evolving population of *Escherichia coli*. *Cold Spring Harb Symp Quant Biol*. **74,** 119-29 (2009)

27. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol*. **11,** R106 (2010).

28. Flores, N. *et al*. Adaptation for fast growth on glucose by differential expression of central carbon metabolism and gal regulon genes in an Escherichia coli strain lacking the phosphoenolpyruvate:carbohydrate phosphotransferase system. *Metab. Eng*. **7,** 70–87 (2005).

29. Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc*. **3,** 1101–1108 (2008).

30. Bligh, E. & Dyer, W. A rapid method of total lipid extraction and purification. *Can J Biochem Physiol* **37,** 911–917 (1959).

31. Liu, K.-S. Preparation of fatty acid methyl esters for gas-chromatographic analysis of lipids in biological materials. *J. Am. Oil Chem. Soc*. **71,** 1179–1187 (1994).

32. Hegreness, M., Shoresh, N., Damian, D., Hartl, D. & Kishony, R. Accelerated evolution of resistance in multidrug environments. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 13977–81 (2008).

33. Barrick, J. E., Kauth, M. R., Strelioff, C. C. & Lenski, R. E. Escherichia coli rpoB mutants have increased evolvability in proportion to their fitness defects. *Mol. Biol. Evol*. **27,** 1338–47 (2010).

34. Cooper, V. S., Schneider, D., Blot, M. & Lenski, R. E. Mechanisms Causing Rapid and Parallel Losses of Ribose Catabolism in Evolving Populations of Escherichia coli B Mechanisms Causing Rapid and Parallel Losses of Ribose Catabolism in Evolving Populations of *Escherichia coli* B. *J. Bact*. **183,** 2834-2841 (2001).

35. Brissette, J. L., Weiner, L. & Ripmaster, T. L. Characterization and sequence of the Escherichia coli stress-induced psp operon. *J. Mol. Biol*. **220,** 35–48 (1991).